

PROTECTION OF A PROGRAM WAITING TO BE EXECUTED IN A MEMORY USED BY A MICROPROCESSOR

PRIORITY CLAIM

5 **[1]** This application claims priority from French patent application No. 03/04149 filed on April 3, 2003, which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

10 **[2]** The present invention relates to the field of micro-processors and more specifically to multitask circuits capable of executing, at least from the user's viewpoint, several programs simultaneously. In practice, a single instruction of a program is processed at each time by a central processing unit (CPU) but several programs or program sections are stored in a RAM by the CPU which transfers lines of the program to be executed to a cache memory assigned thereto.

DISCUSSION OF THE RELATED ART

15 **[3]** **FIG. 1** is a schematic block diagram of an example of a simplified architecture of the type to which the present invention applies. A CPU **1** equipped with a cache memory **2** (CACHE) is connected to one or several buses **3** of communication (data, addresses, control signals) with peripheral elements. Among the peripheral elements, a RAM **4** intended to contain the data (operands) of the programs being processed as well as the code lines of these programs (at least by blocks) is connected to bus **3**. In practice, the programs contained in memory **4** generally originate from a mass storage **5** (MMEN), for example, a hard disk. This mass storage contains the programs likely to be executed by the CPU as well as the data saved when these programs are not being executed. Of course, several mass storages (CDROM, floppy disk, etc.) may be connected to bus **3**.

25 **[4]** To execute several different applications or programs temporarily stored in memory **4**, CPU **1** must have a table of correspondence between so-called virtual addresses of the program, which are independent from its storage location, and so-called physical addresses corresponding to the physical addresses in the memory, for example memory **4**, where the

different program lines are stored. This correspondence table is generally stored in buffers and is generally designated as a TLB (Translation Look Aside Buffer).

[5] FIG. 2 schematically and functionally illustrates a correspondence table **10** and the exchanges between the cache memory or register containing this table and the different elements using it.

[6] On the side of CPU **1**, each time an instruction of a program stored in RAM **4** must be executed, this instruction is called by using virtual address VirtAD of this instruction corresponding to the address contained in the program. This virtual address is converted by table **10** into a physical address PhysAD where this instruction is located in RAM **4**. RAM **4** then provides the corresponding instruction over the bus (**3**, FIG. 1) to the CPU.

[7] If table **10** does not contain the correspondence between the two addresses, the CPU or, more specifically, a calculation program (block **11**, CALC) of its exploitation system calculates a new correspondence line between the virtual address and the physical address, and writes it into correspondence table **10**.

[8] Each time an application contained in RAM **4** must be executed by the CPU, the exploitation system takes over and uses its internal structures to calculate the correspondence table for the involved program.

[9] FIG. 3 illustrates the fact that two different applications APPL**1** and APPL**2** stored in RAM **4** at different physical addresses (PhysAD(k) to PhysAD(m) for the first program and PhysAD(n) to PhysAD(p) for the second program) share identical virtual addresses (VirtAD(1), VirtAD(2), VirtAD(i), etc). Each time a program comes to the foreground, that is, is executed by CPU **1** (possibly after transfer into cache memory **2**), its correspondence table **10** is used to convert its virtual addresses into physical addresses in memory **4**.

[10] For security reasons in the program execution, it is important for two identical virtual addresses of two different applications not to point to the same physical address in memory **4**. Indeed, this enables protecting the program code, as well as the application data, from one another. For example, if critical data (secret quantity, secret result, etc.) of a first application are located at a physical address between addresses k and m in memory **4** while they are assigned a virtual address i, absolutely no other application must be able to access the

corresponding physical address by reusing virtual address *i*. In an architecture of the type in **FIG. 2**, the correspondence table must be reset each time it is switched from one application to another. This imposes recalculating the correspondence table each time an application switches to a foreground execution, that is, one of its instructions must be executed by the CPU.

[11] An obvious disadvantage is that this takes time.

[12] **FIG. 4** schematically illustrates the architecture of a conventional solution avoiding recalculation of the correspondence table each time an application comes to the foreground. This solution uses an additional field of correspondence table **10** which contains, for each line, an identifier ASID of the application in addition to the virtual and physical addresses. This enables, in principle, avoiding association of an application with a physical address corresponding to another application.

[13] To implement this solution, the circuit uses a register **20** (ASIDREG) containing the identifier of the current application. This register is a state register of CPU **1**.

[14] This solution has the advantage, as compared to that of **FIG. 2**, of not requiring resetting or emptying the correspondence table each time the foreground task is changed.

[15] A disadvantage however remains that it is the exploitation system that assigns identifiers ASID to the different applications when an application is loaded into the RAM for execution. The area of RAM **4** assigned to a suspended (background) application accordingly becomes more vulnerable to possible attacks. Indeed, from the moment that the correspondence table is not integrally recalculated when the application comes to the foreground, it is possible to replace the RAM with an emulator, or to modify the signals on bus **3**, to modify an operation code of a background application and insert a pirate instruction (of virus or "Trojan horse" type) therein. Then, when the modified application comes back to the foreground, it can then have access to the physical addresses which have remained present in the correspondence table upon former execution of the modified application (*i.e.*, before it has been modified).

[16] This vulnerability of a background application is increased by the fact that it is not necessary to be synchronous with the CPU to modify the lines of this memory.

[17] A similar problem is posed for mass storage **5** when the program is not integrally loaded into RAM **4** and its execution requires data line calls in mass storage **5**.

SUMMARY OF THE INVENTION

[18] An embodiment of the present invention aims at providing a method for allowing access
5 to a correspondence table which overcomes the disadvantages of known solutions. In particular, an embodiment of the present invention aims at avoiding systematic recalculation of a correspondence table at each foreground loading of an application to be executed, without making it vulnerable when it is in the background in a RAM.

[19] An embodiment of the present invention also aims at providing a solution which is
10 compatible with conventional uses of correspondence tables.

[20] An embodiment of the present invention provides a method for allowing access to a table of address correspondence between a multitask CPU and at least one memory containing several programs, consisting of calculating, on each task change of the CPU, a signature of at least part of the program instruction lines and of checking the conformity of this signature to a
15 signature recorded upon previous execution of the involved program.

[21] According to an embodiment of the present invention, said signature is calculated by the implementation of a Hash function.

[22] According to an embodiment of the present invention, said memory is a RAM in which are loaded program lines from a mass storage.

[23] An embodiment of the present invention also provides a processor of multitask execution of several programs, exploiting a table of correspondence between virtual addresses of the lines of the different programs and physical addresses of these lines in at least one memory, each correspondence being associated with an identifier of the involved program, comprising means for calculating a current signature based on at least part of the program lines in said
20 memory, and means for comparing this signature with the identifier of the program stored in the correspondence table.

[24] According to an embodiment of the present invention, the identity of the signature and of the program identifier allows the CPU to execute the instruction of the involved program.

[25] The foregoing features and advantages will be discussed in detail in the following non-limiting description of specific embodiments of the present invention in connection with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

5 [1] **FIGS. 1 to 4**, previously described, are intended to show the state of the art and the problem to solve; and

[2] **FIG. 5** very schematically shows in the form of blocks an embodiment of a microprocessor architecture implementing the correspondence table access authorization method according to an embodiment of the present invention.

DETAILED DESCRIPTION

10 [3] Same elements have been designated with same reference numerals in the different drawings. For clarity, only those steps of the method and those elements of the processor that are necessary to the understanding of embodiments of the present invention have been shown in the drawings and will be described hereafter. In particular, the calculation of actual
15 correspondence tables has not been detailed, as such calculations can be implemented with conventionally-used calculation tools.

[4] A feature of an embodiment of the present invention is to calculate, at each context change (coming of a new application to the foreground), the identifier of the application by means of an algorithm executing a Hash function or the like calculating a signature of at least part of the
20 code of the application stored in the RAM and/or in the mass storage.

[5] Another feature of an embodiment of the present invention is to check the conformity of this current calculated signature with respect to a previously-calculated reference signature stored in the correspondence table. The reference signature calculation is performed on each calculation of a new correspondence between a virtual address and a physical address. This
25 signature however always remains the same for a same application.

[6] **FIG. 5** schematically illustrates, in the form of blocks, an embodiment of the present invention. The representation of **FIG. 5** mainly shows the functional links between the different elements.

[7] An embodiment of the present invention exploits an architecture of the type previously described in relation with **FIG. 4**, that is, exploiting an identifier ASID of the application or program associated with each correspondence line of table **10** between a virtual address VirtAD and a physical address PhysAD. As previously described, correspondence table **10** is used to convert virtual addresses of an application required by a CPU **1** provided with an application identifier register **20** (ASIDREG), into a physical address PhysAD of a RAM **4** in which is stored the concerned application. The CPU comprises hardware or software means (block **11**, CALC) for having the physical addresses calculated by its exploitation system either at the loading of an application, or when a virtual address is called for the first time.

[8] According to an embodiment of the present invention, each time an application (for example, a program or a sub-program) contained in memory **4** (or in a mass storage not shown) comes to the foreground, that is, at least one of its instructions is executed by CPU **1**, a signature (block **30**, HASH) of at least part of the lines of the program stored in memory **4** is calculated. This signature provides a current application identifier (CURASID) stored in register **20**. The content of register **20** is then compared (block **32**, COMP) with the identifier ASID stored on the line of table **10** corresponding to the involved application which is desired to be used. The result of this comparison enables verifying that the operation code of the application has not been modified in the RAM while this application was in the background. Comparison block **32**, be it hardware or software, provides an authorization or authentication signal AUT to CPU **1** to take the appropriate measures. In practice, the CPU will only execute, or will only transfer into its cache memory for execution, the instructions of the program stored in RAM **4** if comparator **32** has authenticated that the operation code has not been modified since its loading into memory **4**.

[9] Preferably, the signature calculation is performed on a fixed significant portion of the code of the program stored in the RAM. "Fixed portion" means that lines containing data processed by the program, and the content of which is thus likely to change and to modify their signature even when no piracy has occurred, should be avoided. "Significant" means that the larger the number of code lines taken into account in the signature calculation, the more the authentication will be robust in terms of efficiency. As an example, the signature can be calculated taking one line out of ten, one line out of twenty, or one line out of thirty of the operation code.

[10] An advantage of an embodiment of the present invention is that a modification of the operation code of a background program stored in a RAM becomes difficult since its operation code must be modified while respecting the signature, the calculation algorithm of which is in principle unknown. Preferably, to guarantee the system security and for rapidity reasons, the Hash function is executed in hardware fashion in an integrated circuit.

[11] The application being executed necessarily is in the cache memory of the microprocessor which forms an area considered as untouchable for a pirate. Only when the application is pending in the RAM does there exist a risk of piracy.

[12] It should be noted that if the totality of a program is not transferred from the mass storage (5, FIG. 1) to the RAM at the application loading, the signature calculation can exploit program lines still present in the mass storage.

[13] An advantage of an embodiment of the present invention is that, without requiring recalculation of correspondence table 10 on each coming to the foreground of a new application, it prevents an operation code from being pirated when it is in the background in a multitask processing. Of course, the content of the lines of table 10 can be overwritten as it is filled, as conventionally used to be the case in the solution discussed in relation to FIG. 4.

[14] Any conventional algorithm executing a function of Hash type may be used. Among known algorithms, one can mention, for example, the algorithm known as SHA-1 which operates on 512-bit blocks and provides a 160-bit signature as an output. For the application of such an algorithm, the code or code portion of which a signature is desired to be obtained is cut into blocks of 512 bits for which are calculated, each time, five concatenated 32-bit words corresponding to the block signature. In this case, a single word out of the five 32-bit words of the signature may be used and the first words of several signatures calculated on different blocks may be added to obtain the current code CURASID of the application.

[15] The architecture illustrated in and described with reference to FIG. 4 may be part of an electronic system, such as a computer system.

[16] Of course, embodiments of the present invention are likely to have various alterations, modifications, and improvements which will readily occur to those skilled in the art. In particular, the selection of the signature calculation algorithm preferentially depends on the size of the application identifiers used in the correspondence table and on the security level desired for the

system. Further, the selection of the operation code lines to be taken into account in the signature calculation is within the abilities of those skilled in the art based on the functional indications given hereabove. Finally, although a hardware implementation of the signature calculation is preferred, embodiments of the present invention do not exclude a software implementation.

[17] Such alterations, modifications, and improvements are intended to be part of this disclosure, and are intended to be within the spirit and the scope of embodiments of the present invention. Accordingly, the foregoing description is by way of example only and is not intended to be limiting.